ERIC Document Resume	See Instruction Sheet	DATE OF RESUME	DATE RECEIVED

A. PUBLICATION DATA

1. ERIC NUMBER	2. ORIGINATOR CONTROL NO(S).	3. IDENTIFICATION NO.	DO NOT WRITE IN SPACE BELOW
ED001109	N.A.	N.A.	

4. AUTHOR(S):
Action for Boston Community Development, Inc.

5. TITLE OF DOCUMENT ABSTRACTED:
Experimental Developmental Reading Program in the Boston Public Schools: Program and Research Design

6. PUBLICATION TITLE:
N.A.

7. EDITOR(S):
N.A.

8. PLACE OF ORIGIN
Action for Boston
ORG. Comm. Dev., Inc.
ST. 18 Tremont St.
CITY, STATE Boston, Mass.
TELE. EXT.
INST/LOC CODE

9. PUBLISHER OR SOURCE
Action for Boston Comm. Dev., Inc.
Boston, Mass.

10. VOL.	11. PAGES	12. MO./YR.	13. TOTAL PP.
N.A.	N.A.	Dec 64	27

MICROFILM COPY AVAILABLE YES ☐ NO ☐

COPYRIGHT GRANTED YES ☐ NO ☒

COPYRIGHT RELEASE FILED YES ☐ NO ☐

TYPE OF REPORT
PROPOSAL ☒
IN-PROGRESS ☐
FINAL ☐
FOLLOW-UP ☐

OTHER ERIC REFERENCES:

ERIC SATELLITE CODE

B. INDEXING DATA

14. SUBJECT AREA

15. TYPE OF ENDEAVOR
report

16. INDEXING TERMS (culturally deprived, research design, classroom environment, reading Developmental Reading Program, program evaluation, materials, procedures, ABCD disadvantaged youth, general reading, vocabulary, reading consultant, word recognition skills diagnostic testing, reading comprehension, motivation, visual perception skills, programs

17. ABSTRACT (MAX. 250 WORDS) A program to reduce delinquency is presented. The specific objective is to improve reading performance. The program is the result of a cooperative effort of Boston Public Schools and ABCD, a private organization. The program will concentrate on three areas of the city of low socio-economic level. Test results which are described indicate that reading achievement in the schools of these areas is consistently below the city-wide median, and that the pupils tend to fall progressively farther behind. In the program, reading consultants will be assigned to schools in three elementary and three junior high school districts. Increased time will be devoted to reading instruction in elementary schools, and time will be set aside in junior high schools for reading instruction. Where possible, students will be grouped according to ability level, not grade placement.

The main objectives are to increase general reading vocabulary and to improve reading comprehension. To achieve these, the program aims to improve specific skills, attitudes, and classroom conditions. Each of nine factors are considered in detail and the materials and procedures to be used are described. Among them are word recognition and visual perception skills, comprehension, vocabulary development, and critical reading skills. Motivation will be improved, and interest in reading and appreciation of good literature developed, and a stimulating classroom environment provided.

To facilitate grouping and to determine individual skill deficiencies, diagnostic testing is necessary. Enrichment classes for the more able students will be held. Reading consultants act as resource and demonstration personnel with classroom teachers. The research design for evaluation of the program is described. The evaluation should answer: 1) who were exposed to the program, 2) what procedures and methods were used, 3) did those exposed change in the direction of the objectives, 4)can those changes be attributed to the program?

ERIC No.

18. SOURCE OF ABSTRACT

C. FUNDING DATA

19. SUPPORTING ORGANIZATION	20. DIVISION OR DEPT.	21. CONTRACT NO. / PROJECT NO.
	22. TOTAL COST OF PROJECT	
	23. DURATION OF PROJECT	U.S. Department of Health, Education, and Welfare Office of Education

ERIC Form 1

REPRINTED BY PERMISSION
U. S. DEPARTMENT OF HEALTH, EDUCATION AND WELFARE
OFFICE OF EDUCATION

THIS DOCUMENT HAS BEEN PRINTED EXACTLY AS RECEIVED FROM THE PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPINIONS STATED DO NOT NECESSARILY REPRESENT OFFICIAL OFFICE OF EDUCATION POSITION OR POLICY.

ACTION FOR BOSTON COMMUNITY DEVELOPMENT, INC.
and
THE BOSTON PUBLIC SCHOOLS

EXPERIMENTAL DEVELOPMENTAL READING PROGRAM
in
THE BOSTON PUBLIC SCHOOLS

PROGRAM AND RESEARCH DESIGN

December, 1964

Joseph S. Slavet
Executive Director, ABCD

Robert Perlman
Associate Executive Director
Director of Program Development, ABCD

Clarence Sherwood
Associate Executive Director
Director of Research, ABCD

Harold C. Haizlip
Program Supervisor (Education), ABCD

Irene McInnis
Research Project Director, ABCD

Herbert D. Simons
Program Specialist, ABCD

William H. Ohrenberger
Superintendent of Schools
Boston Public Schools

Marguerite Sullivan
Deputy Superintendent of Schools
Boston Public Schools

Catherine M. Maney
Coordinator, Boston Public Schools - ABCD Programs

Agatha Howard
Assistant Coordinator, Boston Public Schools - ABCD Programs

ED001109

FOREWORD

The following pages describe the action and research design of an experimental reading program developed jointly by the Boston School Department and Action for Boston Community Development, Inc. (ABCD). The program is being financed jointly by the School Department and by funds provided by The Ford Foundation to ABCD.

This demonstration reading program and 15 other programs, listed at the end of this Foreword, were planned under a grant from the President's Committee on Juvenile Delinquency and Youth Crime. Taken together these 16 programs comprise the Boston Youth Opportunities Project, which has as its ultimate objective reducing delinquency among males 12 through 16 in a prescribed area in Boston.

Each program in the Project has a specific objective - in this case, improving reading performance. The program described below will test two assumptions: (1) that the procedures and methods of this program will improve reading performance and (2) that if such improvement does occur, there will be less delinquency among those exposed to the program than among a comparable group not exposed to this program.

The planning and implementation of this program represent a collaborative effort between a public school system and an organization, ABCD, concerned broadly with community development. ABCD is a private organization incorporated in 1962 to design new programs in employment, education and social services; to supply partial financing for these programs; and to help public and private institutions to put them into effect.

ABCD was formed in 1961 as the result of the strong conviction on

the part of Mayor John F. Collins, Development Administrator Edward J. Logue, and a group of civic leaders that Boston's comprehensive urban renewal program provided both the opportunity and the necessity for a program focused on the social needs and problems of the city. ABCD was launched with initial support from the Committee of the Permanent Charity Fund, Inc., a Boston foundation, and from United Community Services of Metropolitan Boston, the area's health and welfare council.

ABCD was envisioned as an organization with a limited life span, designed not to replace existing agencies but to generate new approaches to Boston's social problems. Nor was it to be a service agency; rather, it would provide the planning, research, financial and other assistance necessary for public and private organizations to carry out new programs.

From the outset, the Superintendent of Schools and a member of the five-man Boston School Committee have served on the Board of Directors of ABCD, together with a widely representative cross-section of the city's governmental and civic leadership. On a staff level, a Coordinator of Special Projects was appointed by the Boston School Department to facilitate the planning and supervise the implementation of the demonstration programs. A Supervisor of Educational Programs was employed by ABCD to develop the school programs in cooperation with school personnel.

The funding of this and the other educational programs in the Boston Youth Opportunities Project provides that the Boston School Department will assume increasing financial responsibility during the three-year life of the demonstration. Specifically, during the first year the schools will finance approximately 25% of the cost of the program and ABCD will finance

75%; during the second year the sharing of costs will be equal; in the third and last year of the demonstration, the schools will assume 75% of the cost and ABCD 25%. If a careful and scientific evaluation determines that the program, or some part of it, has demonstrated its effectiveness in improving reading performance, it is understood that the school system will then consider, not only the full financing of the program, but its expansion to other schools and parts of the city that may presumably benefit from it.

Boston's public schools reflect the social and economic changes the city has been experiencing for the past two decades. One basic change results from the out-migration of white middle-class families and the in-migration of low-income, predominantly Negro, families. These population shifts, coupled with the educational deprivations of many of the incoming group, now confront Boston and its schools with the task of resolving new and urgent problems.

Although the total population of the city decreased from approximately 800,000 in 1950 to 700,000 in 1960, the public school enrollment remained remarkably stable. There were 92,793 pupils in the public schools in 1962, an increase of about 500 since 1952. The present population includes a larger proportion of school-age children than was the case ten years ago.

The public school-ABCD demonstration programs are concentrated in three areas within the City of Boston: Roxbury, the South End and Charlestown. Roxbury, with a population of 85,000, has been absorbing increasing numbers of Negroes in recent years; it is estimated that they may now

constitute 40-50% of the area's population. The South End, a community of 34,000, contains substantial numbers of Negro, Puerto Rican, Chinese, Syrian and other ethnic groups. Charlestown is an all-white, predominantly working-class community of 20,000.

These are the areas of Boston where large numbers of families are engaged in a basic struggle for survival and in a continuing conflict with economic and environmental problems. A disproportionate number of their children are having difficulties in school. Unless concentrated efforts are made to deal with their problems of learning retardation, failure, premature dropout, truancy, delinquency and parental indifference to the need for education, countless youth in these families will continue to waste their potential. That waste will be a serious loss to the community.

INTRODUCTION

The entire academic curriculum of the school is built on the expectation that the child can read. Not only is the ability to read a necessity in each of the subject areas as a key to the content, but the ability to understand written instructions and explanations is essential throughout the child's future in the world outside of school.

Test results, described below, indicate that reading achievement in a number of Boston schools is consistently below the city-wide median. The test data show, moreover, a tendency for pupils in these schools to fall progressively farther and farther behind the city-wide norms. This evidence that the problems of teaching and learning how to read have not been resolved is the point of departure for this program of experimentation and research in reading.

A description and explanation of the tests given in the Boston schools is presented here as background information. In addition, the test results are given for the city as a whole and for the schools that are included in the demonstration. (See the section on evaluation for an explanation of "experimental" and "control" schools.)

The first reading test given in Boston in the first grade is the Murphy-Durrell Diagnostic Reading Readiness Test. This test yields scores in three areas: auditory, visual and learning rate. To succeed on this test the child must listen and be able to tell whether two words sound alike or different. Next, the child must find in his test booklet the same letter or word that is being held up by the teacher. Finally, the child is asked to learn ten words during the day. These three areas, taken together, are intended to measure the child's readiness to begin

formal reading instruction.

Boston's schools report scores only for the auditory and visual sections of the test. These scores are reported in percentiles measured against nationally-established norms. The 50th percentile is average; 99th is the highest. Normally, the first percentile would indicate a score at the very bottom of a test. However, on the auditory portion of this readiness test, one-third of the population had <u>one item correct out of a maximum of 84</u>. Therefore, one correct answer would place the pupil at the 30th percentile on the nationally-determined scale, the lowest percentile reported for the auditory section. On the visual portion of the test, one correct answer places the child at the third percentile.

In 1962-63 the Boston city-wide median was the 45th percentile on the auditory portion of the test and the 42nd percentile on the visual section of the test. The following table shows the distribution of scores in the first grade in the demonstration schools for 1962-63. The designations "EE1", "EC1", etc. refer to specific schools; the names of the schools are not used in this paper.

	Murphy-Durrell Diagnostic Reading Readiness Test Percentile	
	Auditory	Visual
Experimental Elementary		
EE1	30	26
EE2	30	31
EE3	30	30
Control Elementary		
EC1	42	26
EC2	60	44
EC3	48	36
EC4	30	14
EC5	30	15
EC6	45	32
City-wide	45	42

The range for all Boston schools was the 91st to the 30th percentile on the auditory section and the 88th to the 14th on the visual portion.

In the second grade the children are given the Gates Primary Reading Tests. This instrument includes three separate booklets: Word Recognition (WR), Sentence Recognition (SR) and Paragraph Reading (PR). The tests are intended to be diagnostic rather than measures of speed.

In the Word Recognition booklet the child selects the word out of four which best describes the picture in the same box. The Sentence Reading Test asks the child to read a sentence and then mark a picture which illustrates the meaning of the sentence. Finally, the child reads several paragraphs and marks the picture in accordance with the directions given in the paragraph.

Scores for this and the following tests are reported by grade levels and in school years and months. For example, a second grade pupil would be expected to score 2.1 in October, 2.4 in January and 2.9 in June. The following September the same child would be expected to score 3.0 Ten months are equivalent to one school year.

For the second grade, the city-wide median in 1962-63 on the Gates Reading Tests was 2.6 for Word Recognition, 2.4 for Sentence Reading, and 2.3 for Paragraph Reading.

The second grade test results for the demonstration schools appear in the following table.

	Grade II Gates Primary Reading Test Grade Equivalent		
	WR*	SR*	PR*
Experimental Elementary			
EE1	2.2	2.2	2.2
EE2	2.1	2.2	2.2
EE3	2.2	2.3	2.3

Control Elementary	Grade II Gates Primary Reading Test Grade Equivalent WR*	SR*	PR*
EC1	2.8	2.6	2.4
EC2	2.1	2.1	2.2
EC3	2.4	2.3	2.2
EC4	2.1	2.0	1.7
EC5	2.2	2.2	2.2
EC6	2.4	2.3	2.3
City-wide	2.6	2.4	2.3

*WR=Word Reading SR=Sentence Reading PR=Paragraph Reading

Third grade children are given the Primary Diagnostic Reading Test. The purpose of this test is to help teachers identify those pupils who need help in reading and to point out particular areas in which these pupils need help. The students are tested on knowledge of words out of context, facility in paragraph comprehension, and ability to read material with adequate speed. The child must be able to follow directions, locate material and recall and comprehend what he has read.

On the Gates tests, only one out of 180 median scores (3 scores each for 60 Boston districts) was below a second grade reading level. On the third grade Primary Diagnostic Reading Test, 108 of 180 median scores were below the third grade reading level. The range of scores from the third grade districts varied from 2.0 to 3.9. The city-wide median was 2.8 (vocabulary), 2.9 (rate) and 2.9 (comprehension). Scores in the demonstration schools are given below.

Experimental Elementary	Grade III Primary Diagnostic Reading Test Grade Equivalent Vocabulary	Rate	Comprehension
EE1	2.0	2.7	2.6
EE2	2.0	2.2	2.5
EE3	2.0	2.2	2.5
Control Elementary			
EC1	3.2	3.1	3.1
EC2	2.2	2.6	2.5
EC3	2.6	2.9	2.8
EC4	2.1	2.7	2.6
EC5	2.1	2.2	2.5
EC6	2.2	2.7	2.7
City-wide	2.8	2.9	2.9

The Durrell-Sullivan Intermediate Reading Test is given to fourth grade children. Speed of reading is not a factor in scoring this test. The relevant parts are word meaning and paragraph meaning. The 1962-63 city-wide median on this test was 4.0. There was a difference of one school year and nine months between the lowest district and the highest district. Scores ranged from 3.0 to 4.9. Median test results for the demonstration schools are as follows:

Experimental Elementary	Grade IV Durrell-Sullivan Grade Equivalent-Median
EE1	3.7
EE2	3.5
EE3	3.4
Control Elementary	
EC1	4.4
EC2	3.6
EC3	3.9
EC4	3.5
EC5	3.3
EC6	3.8
City-wide	4.0

No fifth grade reading test is given. Sixth graders are given the Metropolitan Elementary Reading Test, which measures basic skills of word knowledge and reading. This test emphasizes words which have been found to recur frequently in different series of basic readers. Speed of reading is not an important factor in the test.

On the Metropolitan test the children tend to score lower in comprehension than in word knowledge. However, both the lowest and the highest district scores are in the area of word knowledge. In 1962-63, sixth grade children in the school district reporting the highest scores were advanced over the children in the district reporting the lowest scores by three

school years and seven months. The city-wide median for the sixth grade was 6.0 (word knowledge) and 5.5 (reading).

Grade VI

Experimental Elementary	Metropolitan Intermediate Reading Test Grade Equivalent-Median Word Knowledge	Reading
EE1	5.5	5.4
EE2	4.7	4.4
EE3	4.7	4.7
Control Elementary		
EC1	5.3	5.5
EC2	4.9	4.9
EC3	5.3	5.2
EC4	4.9	4.7
EC5	3.6	4.0
EC6	5.5	4.9
City-wide	6.0	5.5

Reading tests are not administered in the seventh grade. The Iowa Silent Reading Test is given to eighth graders. Several types of reading skills are called for on this test. Speed of reading is measured at controlled comprehension levels. The pupil must understand words, sentences and paragraphs. In addition, he is asked to use dictionary and index skills to locate information. Paragraph material and vocabulary questions

include science and social studies areas.

The city-wide eighth grade median in 1962-63 was 7.6. Children in the district reporting the lowest scores were three school years below children in the district reporting highest scores. The district medians varied from 6.4 to 9.4. In the demonstration schools the results were as follows:

Experimental, Junior High	Grade VIII
JE1	6.4
JE2	7.3
JE3	6.5
Control, Junior High	
JC1	7.1
JC2	6.9
JC3	7.3
City-wide	7.5

A review of the foregoing data clearly indicates the growing gap between reading performance in the four experimental schools and the city-wide median. The following table shows the number of months below the city median of the scores in the experimental schools

DIFFERENCE IN MONTHS IN READING SCORES, 1962-63
BETWEEN EXPERIMENTAL SCHOOLS AND CITY-WIDE MEDIAN

	SCHOOL	WR	SR	PR
Grade II	EE1	4	2	1
	2	5	2	1
	3	4	1	0
		VOC	RATE	COMP
Grade III	EE1	8	2	3
	2	8	7	4
	3	8	7	4
		Word and Paragraph Meaning		
Grade IV	EE1	3		
	2	5		
	3	6		
		Word Knowledge	Reading	
Grade VI	EE1	5	1	
	2	13*	11	
	3	13	8	
		Total Reading Score		
Grade VIII	JE1	11		
	2	2		
	3	10		

* One year and three months below city-wide median

The purpose of this experimental reading program is to arrest and reverse this progressive deterioration in reading skill and performance in the experimental schools in this demonstration.

THE PROGRAM

Overview

The experimental reading program will be conducted in three elementary school districts and three junior high school districts. Reading consultants will be assigned to these schools to work as resource and demonstration personnel with classroom teachers.

Increased time will be devoted to reading instruction in the elementary schools, and for the first time in Boston history the experimental junior high schools will have one hour per day set aside for reading instruction. Wherever necessary and feasible, students will be grouped on the basis of reading achievement rather than grade placement. Each of these elements of the program is described in greater detail below.

Objectives of the Program

This demonstration reading program has two main objectives:

1. To increase general reading vocabulary with special emphasis upon:
   a. mathematics;
   b. science; and,
   c. social science.
2. To improve reading comprehension with special emphasis upon:
   a. following directions;
   b. reference skills; and
   c. interpretation of materials

If the demonstration is to be successful, the program must significantly change its target population in these two basic components of reading achievement.

Program Implementation

To achieve these objectives, the program will make use of prescribed procedures and materials to bring about the improvement of specific sub-skills, attitudes and classroom conditions which are assumed to affect reading achievement. These are:

1. word recognition skills;

2. visual perception skills;

3. comprehension subskills;

4. subskills for reading effectively in different subject areas;

5. vocabulary development subskills;

6. critical reading skills;

7. interest in reading and appreciation of good literature;

8. improved motivation to achieve in reading; and,

9. stimulating classroom environment.

Each of these factors will be considered in detail below and the materials and procedures to be used for their achievement will be described.

I. Word Recognition Skills

A. Skills to be taught (with emphasis on phonetic and structural analysis)

1. Continue sequential development on phonetic analysis skills.

a. auditory discrimination

b. visual discrimination

c. fundamental principles concerning

1. consonants and consonant blends
2. vowels
3. homophones
4. syllabication
5. accent

2. Continue study of structural analysis

a. compounds

b. contractions

c. root words - inflected endings

d. prefixes-suffixes

B. Materials and procedures

1. PHONETIC KEYS TO READING will be used on the elementary levels to develop the skills and abilities needed for independent recognition of words encountered in new material read. This text has been used successfully with children of varying backgrounds, and is based on the belief that reading is a spiral process. Skills which are presented in one grade are repeated, reinforced, and revised in succeeding grades at greater levels of difficulty.
2. The junior high school student will use PHONICS WE USE, Books E and F. BE A BETTER READER and ADVANCED SKILLS IN READING as a basis for reviewing and refining the skills taught on the elementary level. Students will receive guidance in applying these skills to material that increases in difficulty as they progress.
3. The S.R.A. Reading Laboratories will be used to strengthen and extend the skills presented in the above texts. The varied materials provided give the child the needed review of the skills taught at his own rate of speed.
4. Class lessons will be presented and reviewed on the overhead projector to stimulate and maintain interest at a high level.

5. SPEECH TO PRINT PHONICS and WORD ANALYSIS CARDS will be used on the elementary and junior high level. Together they represent a complete phonics program covering the initial teaching of phonics and word analysis practice. These materials can be used at all reading levels, including instruction of junior high school students who are on an elementary reading level.

II. Visual Perception Skills

A. Skills to be emphasized:

1. visual memory
2. vocabulary development
3. speed of perception
4. visual discrimination of word forms

B. Materials and procedures:

1. Skill texts and teacher-made exercises for the overhead projector
2. Tach X - Tachistoscope

Teachers will be trained to use the Tach-X in order to increase their students' sharpness and rate of visual discrimination. Visual memory will be developed by having the student write the word after a quick exposure and correct his spelling. Through the selection of film-strips students will be introduced to new vocabulary which will be systematically taught and reviewed. It is expected that the sensitive and creative use of this machine will enhance student motivation for learning.

III. Comprehension Subskills

A. Skills to be reviewed, refined, and developed

1. Silent reading and study skills

a. following directions

b. recalling facts

c. skimming

d. finding the central idea

e. drawing conclusions

f. predicting outcomes

g. outlining

h. summarizing

i. proving statements

j. reference skills

k. speeded reading

B. Materials and procedures

1. S.R.A. Reading Laboratories

The Reading Laboratory will provide each child with an individualized learning opportunity that enables him to develop his comprehension abilities at his own rate of speed. The comprehension exercises will help the student recognize main ideas, perceive the author's tone and purpose, understand relationships and evaluate ideas. Rate Builder exercises are designed to provide practice in speed and to develop concentration. One of the purposes of the exercises is to

accelerate the student's ability to think clearly. The READING LABORATORIES will provide the basis for a systematic, sequential development of the comprehension skills..

2. On the junior high level these skills will be strengthened and reinforced through the use of the E.D.L. REFERENCE SKILLS LIBRARY, BE A BETTER READER, Smith, and ADVANCED SKILLS IN READING - Gainsburg. On the elementary level the READER'S DIGEST READING SKILL BUILDERS will provide the repetition necessary to ensure the acquisition and development of these skills. The E.D.L. REFERENCE SKILLS LIBRARY, GRADES 4,5, and 6 can be used on the elementary level also.
3. Lessons prepared by the reading consultant and the teacher will be used to maintain interest and to bolster self-confidence. All materials will be adapted to meet the needs of the child until mastery is obtained.
. The consultants will develop demonstration lessons using STANDARD TEST LESSONS IN READING to show how one book of test lessons can be used to teach a number of skills.
4. READING FOR MEANING -by Guiler-Coleman, Grades 4-9 will be useful at all levels in teaching important comprehension skills such as: vocabulary, context clues,detailed questions, main ideas, outlining and drawing inferences.
5. The student's speed of reading can be increased through the use of the controlled reader and the S.R.A. Rate Building

exercises. Flexibility of reading rate will be stressed rather than speed for the sake of speed.

IV. Skills for Reading Effectively in Different Subject Areas

A. Skills to be emphasized

1. Basic reading skills common to all subjects.

2. Particular reading skills necessary to comprehend the content of science, social studies, literature, and arithmetic texts.

B. Materials and procedures

Each content subject has its own unique aims, body of ideas, logic, vocabulary and relationships. Each subject area needs its own group of reading techniques. Although there is a considerable amount of overlap among techniques, ability to read well in one subject does not insure competence in another. The ability to read and understand work problems in mathematics must be developed systematically at the elementary level if the child is to achieve success in mathematics. Toward this end the E.D.L. ARITHMETIC STORY PROBLEMS will be used. This material develops the ability to read problems more effectively, to gather and retain facts, and to comprehend and interpret correctly in an arithmetic situation. Programmed material of this type heightens the student's interest, encourages careful, accurate reading of each story problem and improves the child's facility in problem solving. Nila Blanton Smith's BE A BETTER READER gives training in science, social studies, mathematics, and literature at the junior high level. The special vocabularies included in the book

give preference to the words most commonly used in these subjects. Training is provided in estimating answers, deciding the process to be used, reading graphs, adding using formulas in mathematics. In science and social studies students learn to classify facts, read diagrams and graphs, and read pictures to obtain information. The E.D.L. Science and Social Studies Labs 4,5,6,7, and 8 will be used here.

The reading consultants will contribute to the reading program in this area by developing demonstration lessons from the subject matter books that the students presently are using. This type of lesson is expected to be valuable because it transfers the reading skill directly into the content area.

V. Vocabulary Development

A. Development pro[illegible]:

1. Systematic study of all new vocabulary
2. Continuous study of synonyms and antonyms
3. Work with multiple meaning words
4. Study of vocabulary of the content fields

B. Materials and procedures

On the assumption that enriching the child's vocabulary will contribute to his reading improvement, materials have been selected to accomplish this purpose. These materials give special attention not only to synonyms, antonyms, and multiple meaning words, but also to technical words used in various subjects.

The story film strips which are to be used on both levels with the controlled reader provide for vocabulary development with each story. New words are repeated in succeeding stories until the child has mastered them. The S.R.A. READING LABORATORIES provide training to develop the ability to handle multiple meaning words, and to become increasingly familiar with synonyms and antonyms. The READER's DIGEST SKILL BUILDER, on the elementary level, and WORD WEALTH JUNIOR and BE A BETTER READER on the junior high level are planned to give practice work in intensive vocabulary development.

VI. Critical Reading Skills

A. Specific skills to be developed

1. Ability to distinguish between fact and opinion
2. Ability to evaluate opinion and ascertain the validity of the content
3. Ability to challenge fact or opinion
4. Ability to recognize propaganda techniques

B. Materials and procedures

1. Pupils will be encouraged to discuss stories or books read in class and to recognize similarities and differences in stories previously read.
2. THE READER'S DIGEST and BE A BETTER READER will be used to teach the pupil to evaluate the usefulness of any content for his purpose of the moment.

3. Newspapers, magazines, textbooks will provide ample material to develop the pupil's ability to think critically and so to distinguish fact and opinion.

VII. Interest in Reading and an Appreciation of Good Literature.

A. Development of enjoyment through

1. Opportunities to share stories read

2. Informal and stimulating discussion of stories or poems read.

3. Creative activities, such as:

Painting and drawing
Literature mural
Dramatization
Puppets

B. Materials and procedures

1. The S.R.A. PILOT LIBRARIES provide a substantial collection of good literary stories -- adventure, animals, sports, humor, biography -- topics that have a genuine interest for young people.

Pupils, particularly those in the top groups, will be encouraged, with the cooperation of the librarian, to read realistic, informative, and helpful books on human relations.

VIII. Improved Motivation to Achieve in Reading

Development provided by:

1. Pacing the rate of learning to suit the child. Each child will progress according to his ability. The brighter children will progress faster than the slower children. Homogeneous ability

grouping is designed to facilitate this. The use of SRA Reading Laboratories will allow the children to progress at their own rate.

2. The material must be on the child's reading level. All the material read by the pupil must be on his reading level. Work below the student's level is boring and work above his level is frustrating. In either case he will not maintain the proper motivation. Reading material on his reading level will assure proper motivation, success in reading and maximum progress.

3. The material must be divided into several separate skills. Small units of instruction must be used to enable the child to feel that he is reaching a particular goal. Each lesson will have 3 or 4 different activities. Each activity will have a specific goal which the child will comprehend and then either approximate or accomplish.

4. Motivational questions should be asked before reading. The use of these questions motivates the children to read and remember the stories they read.

5. A variety of lessons should be used. Each lesson will have 3 or 4 different activities and these will vary from lesson to lesson.

6. Graphs and Charts. These will be used as a visual representation of the children's progress. The students can see their progress and thus will be further motivated to achieve.

7. Questionnaires on pupil's personal reactions. Through the use of questionnaires the pupil's likes and dislikes will be ascertained and lessons will be developed around their interests.

IX. Stimulating Classroom Environment

A. Use of programmed materials, such as:

SRA READING LABORATORIES
EDL CONTROLLED READER and filmstrips
EDL Tach-S Tachistoscope and filmstrips
EDL REFERENCE SKILLS LIBRARY

B. Use of a variety of new skill texts designed to develop reading skills:

PHONETIC KEYS TO READING
BE A BETTER READER - Smith
ADVANCED SKILLS IN READING - Gainsburg
WORD WEALTH JUNIOR - Miller
READER'S DIGEST READING SKILL BUILDERS
SPEECH TO PRINT PHONICS - Durrell

C. Provision for creative teaching lessons through use of an Overhead Projector to engage attention and to maintain a feeling of excitement about the reading program.

D. Creation of interesting bulletin boards and meaningful progress charts.

Program Operations

This experimental reading program will be conducted in three of Boston's elementary school districts and three junior high schools during 1964-65, 1965-66, and 1966-67. One junior high school and its feeder elementary schools have been designated for a more intensive program than the other schools participating in the demonstration project.

The experimental groups for the first year will include all the fifth and sixth grade students in two K-6 school districts, all the fifth, sixth and seventh grade students in a K-8 school district, and all the seventh grade students in three junior high schools. The program will be expanded to include fourth grade students in the second year and eighth and ninth grade students in the third year of the demonstration period. The following table shows the number of pupils and Reading Consultants in the schools and grades which the Reading Program is designed to serve in the first year:

School District	Type	Grades	No. of Pupils	No. of Reading Consultants
EE1	K-6	V,VI	284	2
EE2	K-6	V,VI	290	2
EE3	K-8	V,VI,VII	354	2
JE1	Jr. High	VII	200	2
JE2	Jr. High	VII	175	1
JE3	Jr. High	VII	265	1

On the elementary level, two reading consultants will be assigned to each of three elementary school districts. Grades five and six in these districts will participate in the program during the first year. In the second year, grade four will be added.

A key feature of the experimental program is the breakdown of the traditional class and grade confines. In June of 1964, all elementary students in grades four and five of the demonstration schools will be given the California Reading Test. This test consists of two parts, Test 1 and

Test 2. Each test has several sub-tests. The test yields a total score on Reading Comprehension, plus sub-scores on Following Direction, Reference Skills and Interpretation of Material.

On the basis of their scores, students will be assigned in the following September to homogeneous reading achievement groups which combine students from grades five and six for instruction in reading daily, throughout the school year. During the second and third years of the demonstration project, the same test will be administered to students in grades three, four and five, and the reading classes will contain students homogeneously grouped from these three grade levels.

Grouping:

In any classroom there is a wide range of reading ability. In addition, there is an even greater variation in reading subskills; two students reading on a given grade level may have very different subskill performances.

In a developmental reading program which includes students reading below and above grade level, cognizance must be taken of this wide range of individual differences. The following principles for grouping will be implemented wherever feasible:

A. The classes should be ungraded. The students should be grouped across grade level. In the K-6 schools the 4th, 5th, and 6th grades should be grouped together. In the Junior High, 7th-9th grades should be grouped together. In the K-8 schools, the 4th, 5th, and 6th grades will be grouped together, and the 7th and 8th will be grouped together.

B. The students will be homogeneously grouped according to reading level.

C. In each class there will be only one reading level represented, with the following exceptions:

1. Students with the lowest reading scores will have classes with as small a range as possible.
2. Students reading above grade level will have classes with a range of more than one year.

D. For students with the lowest reading scores, class size will be 20 or less.

E. Within each class provisions will be made for two or more sub-groups to handle subskill variations.

Diagnostic Testing:

As mentioned above, there is a great variation in reading subskills ability even among students on the same reading level. In order to correct skill area deficiencies, diagnostic testing is necessary. This testing will enable the teacher to ascertain in which areas the students need help and teach to those areas. Also grouping within each class can be done on the basis of this testing. For example, in phonetic analysis if one group of students need vowel sounds and another word analysis practice then the class could be divided into two groups. One group could work on vowel sounds with the teacher, while the other could work independently on word analysis. In addition more than two groups could be employed which would further individualize instruction. The advantages of grouping within each class as opposed to whole class instruction are obvious in the light of the

preceeding discussion.

The following diagnostic tests will be used to facilitate grouping:

1. Word Attack Test. This is a diagnostic test of the subskills involved in phonetic analysis.
2. Visual Memory Test. This is a test of visual retention of whole words and word parts.
3. Instructional Level Test. This is a test of the oral reading level and ascertains what grade level the student can successfully read.
4. Learning Rate. This test determines the number of new words and word parts.

The diagnostic testing will take place twice during the year. The children will be tested in September and January. In September it will be used to determine instructional needs. In January the testing will be used as a basis for changing groups as well as determining instructional needs.

Large blocks of time during each school day will be devoted to specific consultation and instruction in reading. In the elementary schools there will be a forty-minute reading period each morning and a thirty-minute reading period each afternoon, for a total of seventy minutes each day or 350 minutes (5.8 hours) each week. This emphasis upon the development of reading skills will be extended to instruction in the subject matter areas.

In this phase of the program, approximately 880 minutes (or roughly 14 hours) will be available each week for the consultants in elementary schools to observe classroom teaching, to conduct demonstration lessons and to consult with assigned classroom teachers in order to help them further

develop skills in reading as they teach arithmetic, social studies, science, geography, etc. Their goal is to help teachers orient the entire elementary school curriculum toward the teaching and learning of reading skills. The reading consultants will spend four-fifths of their time in these activities.

Each elementary reading consultant will also spend one-fifth of his time conducting one enrichment class in literature for a group of 15 to 20 of the most able students taken from Grades 5 and 6 in the first year and from Grades 4,5, and 6 in the second and third years. The consultants will select these students on the basis of their general school performance and particularly on their success in reading as indicated by their diagnostic reading achievement tests and their I.Q. scores. These students will be at the top of their class in school performance, reading scores, and I.Q.

The enrichment classes will meet once each week for forty minutes. The literature which the group has read independently during the week will be discussed. Each student will read an average of twenty-five books during the school year and will be given opportunities to broaden his reading interests. A basic list of reading selections drawn up by the consultants will be used to stimulate and guide individual reading interests and to complement the regular school curriculum. The selections include novels, plays, biographies, technical journals, current events, etc.

The experimental junior high school program is similar in objectives and methods to the elementary program. In one junior high school there will be a more intensive program. Two reading consultants will be assigned to this school for the first year, with only Grade 7 involved in the program. A third consultant will be assigned in the second year when Grade 8 is added

to the program, and a fourth will be assigned in the third year when Grade 9 is added to the program.

In the other two junior high schools, one consultant will be assigned to each school to carry on the program over the three-year demonstration period. Each of these two consultants will have full responsibility for the experimental program as Grades 8 and 9 join the program in the second and third years respectively.

One class period per day will be devoted specifically to reading instruction at the junior high school level. In June of each year students will be given the California Reading Test and on the basis of their scores will be grouped homogeneously in September for one period of reading instruction daily. In the first year, the homogeneous groups will include only seventh grade students; in the second year, the homogeneous groups will include students from Grades 7 and 8. In the third year, students from Grades 7,8, and 9 will be grouped homogeneously for instruction in reading.

The Role of the Consultant

The reading consultant has a number of functions. His major function is that of bringing about change. In order to do this he must understand how to influence the classroom teacher so that he or she will be willing to change some of the methods and materials that are being used. Some training in consultation is important to accomplish this.

In this program the consultants will be responsible for communicating to the classroom teacher the skills required for effective use of

the materials and equipment. Another major responsibility will be the development of demonstration lessons that introduce new techniques and materials. Some of these will be team learning, ever-pupil response, and self-correcting and self-directing materials.

The consultants will draw up specific schedules for consulting with teachers. These schedules will include: the amount of time spent with each teacher, and the amount of time spent teaching demonstration classes. In addition, the reading consultants will make periodic reports to the Program Specialist at ABCD and the Coordinator of ABCD-School Department Programs in the Boston Public Schools. These reports will include: weekly schedules, problems encountered, materials used and future planning.

Throughout the demonstration period, the role of the elementary and junior high reading consultants will be to act primarily as resource, reinforcement and demonstration agents to classroom teachers. They will spend four-fifths of their time not as teachers of students, but as catalysts who help teachers implement the planned reading program and devise more effective teaching methods and materials to improve reading achievement throughout the curriculum.

They will devote specific periods of time each day as observers and/or instructors of classroom teachers. In these sessions, they also will observe the progress of students and teachers in using the equipment, materials and methods of the demonstration project. Thereafter, the consultants will consult with individual teachers to discuss their classroom observations and to help them improve their performance as teachers. Consultation sessions will be scheduled with each teacher before, during and

after the school day at least once weekly and more often as necessary.

To perform their roles adequately, the reading consultants will participate in a pre-service training program of one month and an in-service program once weekly throughout the demonstration project. In the pre-service training program, the emphasis will be on the skills required for effective use of materials and equipment; on the life and learning difficulties of culturally disadvantaged children; on the sociology of the school; and on the nature and problems of a reading consultant's role in an urban school system. The in-service training program will focus additional attention to these areas, with special emphasis upon individual and group problems as they arise in each consultant's school.

A program as broad and comprehensive as this requires observation and reporting procedures through which problems and progress can be identified and dealt with. Throughout the three-year demonstration period, a member of ABCD's program staff will visit schools and classrooms, will observe and talk with classroom teachers and reading consultants, and will work closely with designated members of the School Department to insure implementation of the demonstration program as planned.

RESEARCH DESIGN

OBJECTIVES OF THE PROGRAM

This program has two sets of objectives: primary and secondary. The primary objectives are:

(1) To increase general reading vocabulary with special emphasis upon growth in:

(a) Mathematics;

(b) Science; and,

(c) Social science.

(2) To improve reading comprehension with special emphasis upon;

(a) Following directions;

(b) Using reference skills; and,

(c) Interpretation of materials.

If the demonstration is to be successful, the program must significantly change its target population in these two components of reading ability which are basic and essential to improved performance in school.

EVALUATION OF THE PROGRAM

By the end of the demonstration period, the evaluation should be able to answer the following questions:

1. Specifically who were exposed to the program (by name, address, age, sex, school performance and other characteristics)?
2. What procedures and methods were used in the program? To what extent were different procedures used with different individuals? To what extent were the same procedures applied in different amounts to different individuals?

3. Did those who were exposed to the program change in the direction called for by the objectives of the program?
4. Can those changes, if they did in fact occur, be attributed to the program?

The Target Population and Exposure to the Program

In order to answer the first question concerning those exposed to the program -- by name, address, age, sex, school performance, reading score, I.Q., and other characteristics -- detailed information will be obtained and kept up to date for all pupils in the experimental schools, including careful records of the extent of their exposure to the program. The school records of all the pupils in the experimental schools will be photographed. Each individual will be assigned a six-digit code number which both facilitates the processing of the data and provides protection for the confidentiality of the information pertaining to each pupil. The data on the school records will be coded and punched into IBM cards and stored for access within the ABCD data retrieval system. In addition to being necessary to identify those exposed to the program information about the characteristics of individuals exposed to the program is essential as a basis for:

a. Selecting comparison groups -- that is, comparable populations not exposed to the programs -- in order to determine whether changes can be attributed to the programs.
b. Determining whether different results are obtained for different types of youth. If a program is effective at all, it may not be equally effective with all types.

c. Determining whether the effects produced are related to the amount of exposure. Therefore, information is needed about each individual's exposure to the program.

Program Procedures

In order to answer the second question concerning the procedures used in the program, information will be obtained and kept up to date on the teaching materials used, the amount of time the consultants spend with the teachers, the nature of the consultations, the selections the teachers make from the suggestions offered by the consultants, the frequency with which they use the materials and procedures in the classroom, etc. If the program is determined to be successful, this information will be needed to repeat the procedures in a continued or expanded program. Moreover, a detailed record of how the program was conducted will almost certainly be more specific and detailed than the original program. If a program is unsuccessful, it may be due to the fact that the original design was not carried out. Rather than reject a potentially useful program, this type of evaluation can be used to determine whether the original design was implemented.

The Measurement of Desired Changes

To determine whether the pupils exposed to the program changed as desired, instruments are required to measure changes in the variables specified in the objectives of the program. Those instruments must be administered to the pupils exposed to the program, at least before and after exposure.

The Metropolitan Elementary Reading Test has been administered by the Boston School Department to all the pupils in the experimental elementary schools and grades. The results of these tests will be available to the ABCD School Project Research unit for evaluation of the reading program. The results of other tests -- reading, achievement, and IQ -- administered to these pupils will also be available from the cumulative school records referred to above.

In addition, ABCD research personnel will administer the California Reading Test and the California Test of Mental Maturity to all the pupils in the fifth, sixth and seventh grades in the experimental schools and in the control schools[1] in the first year of the demonstration, in the fourth, fifth, sixth, seventh and eighth grades in the second year, and in the fourth, fifth, sixth, seventh, eighth and ninth grades in the third year.

The California Reading Tests -- for both the elementary and junior high grades -- consist of two main parts, Test 1 and Test 2. Each test has several sub-tests. Test 1, a Reading Vocabulary Test, is composed of fifty items. In addition to the total score on Reading Vocabulary, sub-scores on Mathematics Vocabulary, Science Vocabulary, Social Science Vocabulary and General Vocabulary, are also obtainable. Test 2, a Reading Comprehension Test, is composed of seventy items. In addition to the total score on Reading Comprehension, sub-scores on Following Directions, Reference Skills, and Interpretation of Material are also obtainable.

---

[1] Control schools are discussed in a later section of this design.

As previously described, the primary objectives for this program are:

(1) to increase reading vocabulary; and,

(2) to improve reading comprehension

Primary Objective #1 -- improved reading vocabulary -- to be measured in terms of changes of scores (Post-Test Scores compared with Pre-Test Scores) on on Test #1 of the California Reading Test.

Primary Objective #2 -- improved reading comprehension -- to be measured in terms of changes in scores (Post-Test Scores compared with Pre-Test Scores) on Test #2 of the California Reading Test.

Secondary objectives would be of great program and research interest, if objective measures could be obtained. The assumptions underlying the need for improving reading skills seem to be:

(1) If reading skill is improved there will be an increased interest in reading; and,

(2) If interest is increased there will be increased motivation to achieve in other academic areas.

Thus an effort will be made to measure changes in interest and motivation; but due to limits set by time, staff and size of sample, these objectives will be low on the priority list.

Test reliability coefficients are available from the California Test Bureau (publishers of the California Reading Test) only for Test 1 and Test2, not for the sub-tests. The evaluation plan includes the scoring of these tests by coefficients and also provide coefficients for the sub-tests.

This testing procedure will begin in January, 1964, and be repeated in June, 1965 and June, 1966. Because of the extension of the program into

the fourth, eighth and ninth grades in the next two years of the demonstration, the future testing populations will be larger.

Measuring the Effects of the Program

In order to answer the fourth question listed previously -- whether changes, if they occur, can be attributed to the program -- control groups are required to provide an estimate of what changes would have occurred had there been no intervention. Since the reading program involves the total curriculum of each of the experimental schools, the random allocation of children into experimental and control groups is not feasible.

Since the control of variables possibly related to the outcome variable -- reading ability-- cannot be handled by random allocation to "treatments" procedure, some form of statistical control must be applied to populations already allocated.

There are a number of alternative procedures available:

(1) Matching the experimental schools with selected "control" schools -- schools with comparable student populations which would not be exposed to the experimental reading programs. If this method were used alone, it would probably not produce sufficiently similar populations of control pupils. The experimental schools were selected on the basis of greatest need for the program. Although there are several schools which undoubtedly contain many students who are similar in reading characteristics to the experimental students, the comparability of their total populations to the experimental school populations is highly questionable. In addition, all these schools, experimental and control

are characterized by high population turnover. Even if the populations of the experimental and control schools were sufficiently similar at the beginning of the demonstration period, they might be very different by the end of the demonstration period.

(2) Matching the experimental school populations, or selected samples of those populations, with control group populations on a frequency distribution basis. Samples of pupils in non-experimental schools would be selected so that their distributions of reading scores, I.Q.'s and several other variables would match the distributions of these same variables for the experimental schools.

The frequency distribution method is feasible, but it is believed that the method selected for this design is less vulnerable to the effects of high rates of selective population change between experimental and control groups and allows for statements to be made about all the types of individuals exposed to the program, not just those who are similar in characteristics to the available controls as does the frequency distribution matching method.

(3) Matching pupils in the experimental schools on an individual - matching basis - by selecting pupils from non-experimental schools who match them on the basis of the following variables: age, sex, I.Q. and reading scores. In order to control the school context within which the pupils are situated as much as possible, control cases would be selected only from the schools which are most similar to the experimental schools. Once the matched experimental and control groups have been constructed, further tests of the comparability of the groups will be made by testing

for significant differences in other variables -- broken homes, number of brothers and sisters, etc.

The major advantages of the individual matching method are:

(a) it will yield sufficiently comparable populations of experimental and control pupils if the correct variables are chosen; and,

(b) it makes no assumptions about the nature of the relationship -- linear or curvilinear -- between the reading pre-test scores and the reading post-test scores.

Its major disadvantage is that in the analysis of the program on reading ability, information concerning the interaction effects between the pre-test reading scores and the remaining independent variables would be lost.

The procedure adopted for measuring the effects of the program is actually a modified version of the three methods outlined above, combined in such a way that a powerful statistical device known as the analysis of co-variance can be applied to the data produced by the action-research demonstration.

The procedures which have been adopted for this research design assume that a linear relationship between the reading pre-test scores and post-test scores holds true for the total population, experimental and control pupils combined. It assumes that although the program may have a differential impact on different types of pupils, the pre-test reading score is not a relevant factor in the typology. Whatever impact the reading program may have, it applies generally throughout the range

of pre-test scores. Although this assumption appears to be the "best bet", it may turn out to be false. Therefore, this linearity assumption will be tested as part of the total analysis of the findings and a set of individual matching procedures, which does not make this assumption, will be described in this design and held in reserve to be used should the findings require that the assumption of linearity be rejected.

The procedure adopted for the measurement of the effect of the Reading Demonstration is described below, including the principles upon which it is based and the steps which will be taken to implement it.

The school records of all the pupils in the control schools will be photographed, making the data on the cumulative school records available to the ABCD research unit for matching and other evaluation purposes. In addition, the California Reading Test and the California Test of Mental Maturity will be administered to all pupils in the control schools. Six elementary schools and three junior high schools have been chosen as the control schools. The following is a comparison of the population sizes of the experimental and control schools.

Estimated Number of Pupils in Experimental and Control Elementary Schools for First Year
5th and 6th Grades Only

Experimental Elementary	Number of Pupils	Control Elementary	Number of Pupils
EE1	290	EC1	301
EE2	300	EC6	227
EE3	260	EC3	159
		EC2	212
		EC5	103
		EC4	240
	850		1,242

Estimated Number of Pupils in Experimental and Control Schools for First Year Junior High School Demonstration Reading Program
7th Grade Only

Experimental Junior High	Number of Pupils	Control Junior High	Number of Pupils
JE1	200	JC2	231
JE3	265	JC3	330
JE2	175	JC1	429
JE4*	104		980
	744		

* Seventh grade of the EE3 Elementary School

Steps in the Analysis

Separate analyses will be made for each of the grades in which the experimental program is being conducted in 1964 - 5th, 6th, and 7th.* The analyses will be made in each succeeding year.

Separate analyses will be made for each of the Reading Tests -- the California Reading Test #1 (Verbal Ability) and California Reading Test #2 (Comprehension). Additional analyses using the same procedures will be made separately of the three sub-scores of the Reading Comprehension Test --- "Following Directions", "Use of Reference Materials", and "Interpretation of Materials."

There will, therefore, be six separate primary analyses -- one for each of the two reading tests in each of the three grades, and three secondary analyses for each of the three sub-tests. Each analysis will also include the following steps:

1. Five variables will be involved: sex, chronological age, mental age, pre-test reading score and post-test reading score.

* Students not living during any period of data analysis in the target area will be eliminated from the study.

2. Sex will be controlled in all analyses.

3. The chronological age distributions and the mental age distributions will each be treated separately. The span for chronological age will be dichotomized, while the span for mental age will be divided into sextiles. One cell of the chronological age dichotomy will contain pupils of an age appropriate to the grade being studied, while the other cell will contain pupils who have been held back. For each of the two analyses, males and females and experimentals and controls will all be combined into one distribution.

4. Two separate analyses of co-variance will be made, controlling for chronological age and controlling for mental age.

Following is a sample cell arrangement for the three-way analysis of fifth graders, controlling for sex, mental age, and treatment:

Fifth Grade

Mental Age Sextiles	Experimentals		Controls	
	Male	Female	Male	Female
1				
2				
3		(adjusted post-test reading scores)		
4				
5				
6				

Alternatively (depending on the number of subjects available), a four-way analysis may be done, which would control for all four variables simultaneously; treatment, sex, mental age, and chronological age. In each model (3-way or 4-way) the criterion (post-test verbal or comprehension score) will be adjusted for the appropriate pre-test measure.

5. For each analysis, cases will be placed in the appropriate cell in accordance with their "measurements" on each of the variables -- experimental or control, sex and mental age or chronological age.

6. It is expected that a considerable proportion of the experimental and control groups will be lost to the demonstration through population movement. The co-variance analysis for measurement of effects will be made at the end of the demonstration period utilizing those subjects who remain and for whom both pre-test and post-test scores have been obtained.

7. Each row being considered separately -- refer back to the cell arrangement above -- the numbers of cases in each cell will, by random selection, be proportionalized across rows and down columns.

8. When the remaining cases are so arranged, the hypothesis will be tested (by analysis of co-variance) that there are no differences among post-test reading score means across rows and among columns that cannot be accounted for by differences in reading pre-test scores.

Specifically these hypotheses may be outlined as follows. Considering a four-way design they would be --

<u>Main Effects</u>

a. The treatment has no effect on criterion change

b. Sex has no effect

c. Mental age has no effect

d. Chronological age has no effect

<u>Treatment Interactions</u>

e. The treatment does not appear more effective with one sex than the other.

f. The effect of treatment does not depend on mental age.

g. The effect of treatment does not depend on chronological age.

h. The effect of treatment does not depend on a combination of sex and mental age.

i. The effect of treatment does not depend on a combination of of sex and chronological age.

j. The effect of treatment does not depend on a combination of mental age and chronological age.

k. The effect of treatment does not depend on a particular combination of sex, mental age, and chronological age.

The basic method to be used is as follows:

Regressive equations -- predicting post-test scores from pre-test scores -- will be computed for the total populations in the analysis, for each column and for each cell.

The post-test predicted scores -- computed from a best fitting regression line using the actual pre- and post-test scores of the subjects in the analysis -- will be divided into two components: the constant "a", which is the amount which must be added to or subtracted from each Y score to predict Y; and the coefficient of regression "b", which is the ratio or rate of change of Y for each unit change of X; "b" therefore represents the slope of the regression line of Y on X.

9. All the post-test scores in each analysis will be adjusted on the basis of the regression coefficient produced in the analysis of covariance computation, by subtracting the amount that the score would be expected to deviate from the population post-test mean. This procedure

can be illustrated as follows:

Individual	Pre-Test Score	Population Regression Equation	Predicted Post-Test Score	Actual Post-Test Score	Population Post-Test Mean	Adjusted Score
#1	X	Y=a+bX	Y^1	Y	$\bar{Y}$	$Y-(Y^1-\bar{Y})$

The Adjusted Score represents the individual's post-test score after the influence of his pre-test score has been statistically removed.

10. For each of the six analyses, the hypothesis that there are no differences in adjusted scores among columns and among cells that cannot be accounted for by sampling error will be tested by analysis of variance. If differences occur, the probability of such differences occurring by chance will be computed and reported. On the basis of the computed probability, recommendations will be made concerning the wisdom of expanding, continuing, modifying or discontinuing the program.

In addition to providing a basis for a decision concerning the effects of the program, the analysis of co-variance and analysis of variance techniques will provide valuable information concerning the differential response to the program of the various categories of subjects as classified for the analyses. A comparison of the relative differences between adjusted scores for comparable experimental and control group cells may reveal that the program had its greatest effects on girls of high mental age, for example, or that although it had an over-all effect on the population, its effect on older children was negligible. This latter finding might be generally true for all three populations or for only the older fifth grade children, for example.

In any event, such findings will be valuable in allocating future effort, both in terms of the possible extended use of the demonstration program and in the location of sub-populations for whom further experimentation is indicated.

Reserve Matching Procedure

As previously mentioned, the analysis of co-variance techniques as applied in this design assumes a linear relationship between the reading pre-test and post-test scores. Should the evidence require the rejection of this assumption, the following procedures will be used, to provide control groups for the measurement of the effects of the program.

The basic approach is to match children in the experimental schools with children from the selected control schools. The major differences in these procedures as compared with the analyses of co-variance is that (a) the pupils will be matched on a pre-test reading score, in addition to those variables--grade, sex, mental age and chronological age -- which will be used as control factors in the analyses of co-variance; and (b) the statistical devices to determine the significance of the differences between experimental and control group post-test scores will be different.

Matching will be done at the end of the (or each) demonstration period, using experimental and control cases who have remained in the population during the defined period and will be done on a "cell membership" rather than exact score basis. Matching on the basis of exact scores involves at least two difficulties:

1. The loss of cases, particularly at the extreme ends of the distributions, is likely to be severe; and,

2. The procedure assumes a stability of the scores that is undoubtedly unwarranted. If the entire sample were re-tested, the resulting pairings of experimental and control pupils from the second matching might be entirely different from those of the first matching; and the differences would not be due to chance factors but to factors relating to the unreliability of the test.

The basic procedure which appears to be indicated is that the matching should be based on relative position in the distribution of scores, using a criterion which though arbitrary can be repeated with the same sample, or elsewhere, independently of the specific array of scores.

The notion of relative position in the distribution of scores implies the idea (unless exact scores are used) of a segment of the score continum. The two basic ways of segmenting the score continum of a distribution are:

a. Through the use of some form of centile intervals -- percentiles, deciles, quartiles, etc. Centiles divide the distribution into segments with equal numbers of cases but, except in the case of rectangular distributions, different ranges of scores form the boundaries of the segments, and,

b. Through the use of some of standard score -- Z-Score, T-Score, C-Score, Stanine Score, etc., Stanine Scores divide the distribution into segments with different frequencies but equal ranges of score boundaries.

The significance for matching procedures of the differences between these two ways of segmenting a distribution is that if centile units are used matching will be more precise (a narrower range of scores will be

eligible for matching) in the middle of the distribution and less precise at the extremes of the distribution and in fact if the distribution is normal or approximately normal (bell shaped when graphed) the score range of the boundaries of the centile interval will be a direct function of that interval's relationship to the median -- the closer to the median of the distribution, the narrower the score range of its boundaries. If standard scores are used (standard scores are computed by dividing the difference between a score and the mean of the distribution by the standard deviation of the distribution), matching will be equally precise all along the continum, but the number of cases in each segment further and further away from the mean in each direction will decrease. In fact if the distribution is normal, the relative frequencies in each segment are directly predictable from the formula for the normal curve giving the area proportions under the curve for each segment. If random pairings are made within standard score segments, matching is likely to be less precise in the segments toward the center of the distribution than in those toward the tails relative to potential precision of matching. On the other hand, if nearest exact scores are matched within segments, precision is again likely to be greater in the center segments as compared with the tail segments.

The principle should be adopted that procedures should be devised which protect against two types of errors of inference:

a. Concluding that the program did not have an effect when it in fact did; and,

b. Concluding that the program did have an effect when it in fact did not.

We also want to protect against the possibility that the program may appear to have had an effect on those types of individuals whom we matched least precisely.

The best way to provide these protections is to construct a number of matched samples. This design will therefore provide a series of matched samples on which inferences concerning the effects of the program may be based.

The following matching procedures will be used:

1. Each procedure described below will be performed separately for the elementary and junior-high school populations.

2. The experimental and control cases will be sorted by sex. All matching and analyses will be performed separately for boys and girls.

3. For each matching variable -- age, I.Q. and Reading Score -- the scores of the experimental and control pupils will be combined and arrayed in rank order. (Each of these arrays provides the basis for the subsequent matching procedures).

4. Gentile Matching -- The distributions of each matching variable will be divided into centile ranges. Ten per cent intervals will be used. Each ranked array of scores will be divided into ten intervals with equal frequencies within each segment. Each of these segments will be assigned a score corresponding to its centile range -- the interval 90th centile given the score 10, the interval 80th to 89th centile a score of 9, etc. For each of the variables, each individual with a score falling within one of the centile intervals will be assigned a score corresponding to that interval. For example, an individual with a score falling in the 70th to 79th centile

interval will be assigned a score of 8.

Using these scores, a series of matched samples will be constructed:

a. Matched pairs having identical interval scores on each pair of variables - age, and I.Q., Age and Reading Score, and I.Q. and Reading Score. For example:

AGE INTERVAL SCORE

	1	2	3	4	5	6	7	8	9	10
1										
2									x	
3										
4						x^1				
5										
6								x^{11}		
7										
8										
9										
10										

The three cells chosen for the example are:

(1) those with I.Q. interval scores of 2 and age interval scores of 9

(2) those with I.Q. interval scores of 4 and age interval scores of 6

(3) those with I.Q. interval scores of 6 and age interval scores of 8

To continue the example, there will be n, n', and n'' individuals in each of the cells, respectively. A proportion, p, p' and p'' of n, n', and n'', respectively, will be experimental cases. The extent to which p, p' and p'' tend to be similar or different from cell to cell will depend upon the distribution of the scores of experimental pupils and the control pupils within the combined distributions for each variable. To complete the examples and description of the matching procedure, let us assume that in cell X, n equals 10 and p equals .5. The five experimental pupils and 5 control pupils with identical interval scores on both variables would be randomly paired. If in cell x', n equals 15 and p' equals .4, then 6 of the control pupils will be randomly

selected and paired with the six experimental pupils; 3 potential control cases would be lost for this matching operation. If in cell X'', n equals 40 and p'' equals .75, 10 experimental pupils will be randomly selected and paired with the 10 control pupils; 20 experimental pupils would be lost for this matching operation.

It should be pointed out again, that in the use of centile ranges, the variability of the scores within each segment is less the nearer the segment is to the center of the distribution. Therefore, the variance of the differences between the raw reading test scores of the paired cases in the center segments will be less the variance of the difference between the raw reading test scores of the paired cases in the more extreme segments. If separate analyses are made for significant differences between the post-test reading scores of the matched pairs, experimentals and controls, within centile segments, the smaller variances in the central segments mean that smaller differences will be statistically significant than in the extreme segments where the variances will be greater, since the variances of the distributions enter into the formula for testing for statistical significance. However, with this procedure, fewer of the cases with relatively extreme scores should be lost in the matching.

b. This same general procedure will be used to construct a three variable -- Age, I.Q., and Reading Score -- matching. Here the matched individuals will have the same interval score on each of the three variables. Loss of cases will be greater using this procedure. A series of three variable matched examples can be constructed, however:

(1) using the same interval score boundaries as for the two variable matching, and;

(2) one or more additional matched samples using increasingly broader interval score boundaries.

5. Standard Scores Matching -- The distribution of each matching variable will be divided into standard score ranges by constructing Stanine Scale Scores as follows (working from the ranked array of scores for each individual separately):

the highest	3.7%	of the scores will be assigned a score of 9
the next highest	6.6%	of the scores will be assigned a score of 8
the next highest	12.1%	" " " " " " " " 7
the next highest	17.4%	" " " " " " " " 6
the next highest	19.8%	" " " " " " " " 5
the next highest	17.4%	" " " " " " " " 4
the next highest	12.1%	" " " " " " " " 3
the next highest	6.6%	" " " " " " " " 2
the next highest	3.7%	" " " " " " " " 1

These Stanine Scale Scores are actually C Scale Scores, an eleven-category standardized scale, collapsed to nine categories by combining the two categories at each end.[1]

If the distribution is normal or nearly normal, each of the Stanine Scale Score intervals should include approximately the same number of raw score units within its boundaries -- about one-half the value of the standard deviation for the distribution.

Individuals will be assigned Stanine Scores corresponding to the Stanine Interval within which their raw score falls. Once these Stanine Scores are assigned to each individual for matching variable, the matching

---

[1] CF. J.P. Guilford, Fundamental Statistics in Psychology and Education, N.Y.: McGraw Hill Book Co., 1950, pp. 302-306

Also, J.G. Peatman, Introduction to Applied Statistics, N.Y.: Harper, Row Publishers, 1963, pp. 69-70.

procedures will be identical as for the Centile Matching Procedures: a series of three paired-variable matched samples and a three variable-matched sample.

With standard score-matching procedures, the variance within each Scale Interval should not vary significantly. However, loss of cases will undoubtedly be more severe in the intervals at either end of the distribution since there will be increasingly fewer cases in each interval proceeding away from the central interval in both directions. Even if the distributions of scores for both the experimental and control pupils have approximately the same shape, the distributions may not be congruent on the same abscissa. For example, all those pupils with Stanine Scores of 9 may be potential control cases and all those pupils with Stanine Scores of 1 may be potential experimental cases. If so, all the cases in both intervals would be lost for matching purposes.

As previously noted each of the procedures appears to have a different flaw and therefore both will be used and the results compared. The Centile Matching involves variable precision of matching from interval to interval but promises more equal retention of cases for matching purposes. The Standard Score Matching Procedure involves relatively equal precision of matching from interval to interval but promises variable loss of cases from interval to interval.

The Criteria of Effect in the Matching Method

If the matching procedures are used, the effectiveness of the program will be measured by a comparison of the reading scores and the components of

those scores for each of the objectives of the program. The scores of experimental pupils will be compared with those of the matched control pupils.

The general definition of the effects of the program is in terms of differences between the experimental and the control pupils in the increments to their Reading Scores, comparing the pre-test and post-test scores of the matched pairs.

Difference Score = Experimental Score minus Control Score.

For each of the matching procedures described above, the algebraic sum of the difference in reading scores on the pre-test between the experimental pupils and the control pupils should be approximately zero. If it is not, the matching procedure would be deemed faulty.

Assuming the matched samples survive this test, the null hypothesis to be tested concerning differences between the post-test reading scores of the experimental pupils and the matched control pupils is that there is no difference.

Each matched pair in each matched sample will be treated as a single case and a pre-test difference score and a post-test difference score will be computed. The score of the control case in the matched pair will be subtracted from that of the experimental case; therefore, when the experimental case's score is larger the difference will be a plus; when the control case's score is larger the difference will be minus.

In order to infer that the program had an effect on the reading scores of the experimental pupils, the mean of the post-test difference scores

should be positive (rather than negative) and significantly greater than zero. If differences occur, the probability of such differences occurring by chance will be computed. Judgments as to the wisdom of expanding, continuing, modifying, or discontinuing the program will be made on the basis of these computed probabilities.

The quality of the inferences that will be made about the effects of the program depend primarily upon:

a. the matched samples being truly matched samples; and,

b. the differences on the post-tests being real and significant.

In each case the null hypothesis that the mean difference equals zero is tested. There are four possible relationships between the acceptance and rejection of true and false null hypotheses:

a. the null hypothesis is true and is accepted

b. the null hypothesis is true but is rejected

c. the null hypothesis is false but is accepted

d. the null hypothesis is false and is rejected

Decisions "a" and "d" are correct; "b" and "c" are incorrect. Decision "b" -- the rejection of a true null hypothesis -- is called Type I error or "alpha" error in statistical inference. Decision "c" -- the acceptance of a false null hypothesis -- is called Type II error or "beta" error in statistical inference.

In constructing the matched samples, Type II error must be guarded against. That is, maximum precaution must be taken against inferring that the experimental pupils and control pupils are alike (that there is no difference) on the pre-test reading scores if they are in fact not alike.

In testing for significance or difference between the experimental and control pupils on the post-test, Type I error must be guarded against. Precaution must be taken against inferring that the experimental pupils and control pupils are significantly different on the post-test if they are in fact alike.

Since, as previously noted, the population under study is characterized by a high rate of residential mobility, the individually matched-pair procedures will allow for a continual correction of both the experimental and control groups as they are affected by movements within the population. There are two basic types of population movement which will affect the research design:

1. An experimental pupil changes; and,

2. A control pupil changes schools.

Within each of these there are several sub-categories:

1a An experimental pupil changes to another experimental school. No changes would be required in the membership of the experimental and control groups.

1b An experimental pupil changes to a control school. Such cases -- each experimental pupil and his matched control pupil -- would be removed from the primary experimental and control groups and studied as special populations. At least one basic question could be explored through study of this population if it attains sufficient size. In those cases where experimental cases show improvement in reading ability -- are relative to their matched controls -- while they are

exposed to the program, do they continue to show increased relative improvement while in the non-experimental school?

1c An experimental pupil changes to a non-experimental, non-control school. Such cases -- each experimental pupil and his matched control pupil -- would be removed from the primary experimental and control groups. This population would also provide a basis for special study similar to group "1b" above. In addition, since the non-experimental, non-control schools are likely to be better schools (as judged by the performance of their population) than the experimental schools, another basic question could be explored. In those cases where experimental pupils do not show improvement in reading ability -- relative to their matched controls -- while they are exposed to the program, do they show relative improvement while in a non-experimental but presumably better school?

1d An experimental pupil moves out of Boston, dies or for other reason ceases to attend school. These cases, both the experimental and the matched control pupil, would be removed from the experimental and control groups.

2a A control pupil changes to a non-experimental, non-control school. Such cases -- each control pupil and his matched experimental pupil -- would be removed from the primary experimental and control groups. This population would also provide a basis for special study. A study could be made of

exposure to a new school -- usually a better school -- as compared with the effects on the experimental pupil of exposure to the experimental reading program.

2d A control pupil moves out of Boston, dies or for other reason ceases to attend school. These cases, both the control case and his matched experimental pupil, would be removed from the primary experimental groups.